IDS 576: Assignment 4

Turn in solutions as a pdf+.ipynb on Blackboard.

Note: Answer the following questions concisely, in complete sentences and with full clarity. If in doubt, ask classmates using the forum. Across group collaboration is not allowed. Always cite all your sources.

1 Factorization I (5pt)

Let A be a random variable (RV) with support $\{0, 1, 2\}$. Similarly, let B, C and D be random variables with supports $\{0, 1\}$, $\{1, 2, 3\}$ and $\{10, 20\}$.

1. Write down the joint distribution of A, B, C and D in a factored form. How many numbers (parameters) are needed to fully specify this joint distribution? Write down all factorizations that are possible for $P(A, B, C)$ and $P(A, B)$.
2. If we know that $P(A|B, C, D) = P(A|B)$ and $P(C|D) = P(C)$, then what is the number of parameters needed to specify the join distribution $P(A, B, C, D)$?
3. If we know that $P(B, C, D)$ respects DAG $B \rightarrow C \rightarrow D$, then does it imply $P(C|B, D) = P(C|B)$?
4. If we know that $P(A, B, C, D) = P(A)P(B)P(C)P(D)$, how many parameters are needed to represent the joint distribution?

2 Factorization II (5pt)

Let X_i for $i = 1, 2, 3$ be an indicator random variable for the event that a coin toss comes up heads (happening with some probability p). Assume X_i are independent. Let $Z_4 = X_1 \oplus X_2$ and $Z_5 = X_2 \oplus X_3$ where $\oplus$ denotes the XOR (exclusive OR, see https://en.wikipedia.org/wiki/Exclusive_or) operation.

1. Show the computations of the following: $P(X_2, X_3|Z_5 = 0)$ and $P(X_2, X_3|Z_5 = 1)$.
2. Draw a DPGM and write down the corresponding conditional probability tables. What independence relationships are captured by the DPGM?
3. Draw a UPGM and write down the corresponding factors/potentials. What independence relationships are captured by the UPGM?
4. Under what conditions on p would $Z_5 \perp X_3$ and $Z_4 \perp X_1$? Are these independences captured by the above two graphs? Explain.

3 D-separation (5pt)

Let $A = \{X_2\}, B = \{X_3, X_5\}$ and $C = \{X_1, X_6\}$. Let the DPGM be as shown in Figure 1. Is $A \perp B \mid C$? Justify your answer.

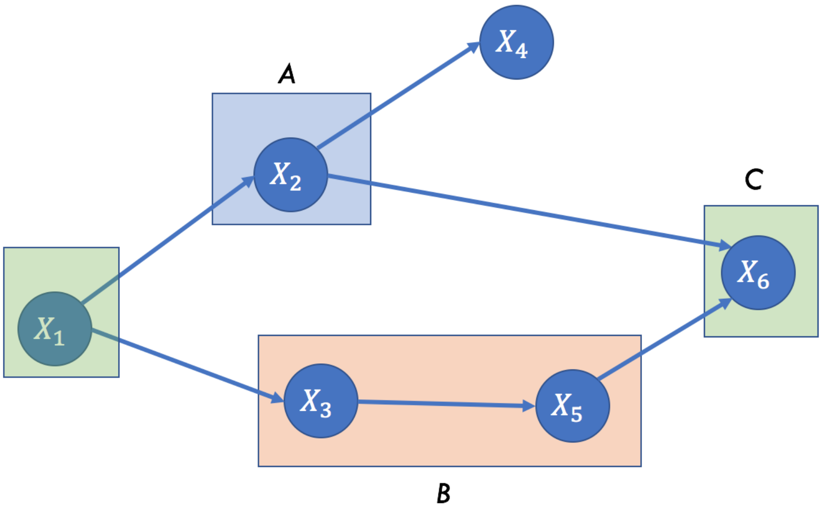

Figure 1: DPGM for Question 3.

4 Inference on DPGM (10pt)

Consider the DPGM in Figure 2 that represents a maintenance sensor network for a machine that manufactures two goods. Each G_i represents the health of a component in a machine. $G_i = 1$ if the component is running and $G_i = 2$ if the component failed. G_1 is the common component needed for both goods whereas G_2 and G_3 are specific the goods. Also, G_2 and G_3 can be influenced by the failure of G_1. X_i is a continuous random variable that measures the quantity of each good type produced by the machine, which is high if the component is running and low if it is not. The conditional probability distributions are:

$$P(G_1) = [1/2, 1/2]$$
$$P(G_i = G_1|G_1) = 0.8, \ i \in \{2, 3\}$$
$$P(X_i|G_i = 1) = \mathcal{N}(X_i|\mu = 100, \sigma^2 = 10)$$
$$P(X_i|G_i = 2) = \mathcal{N}(X_i|\mu = 10, \sigma^2 = 20)$$

1. If we observe $X_2 = 100$, what is the posterior belief on G_1. That is, compute $P(G_1|X_2 = 100)$, and show your work.
2. If both X_2 and X_3 are observed, then what is $P(G_1|X_2, X_3)$? In particular, what are the values when the observations are: (a) $X_2 = 100$ and $X_3 = 100$, (b) $X_2 = 10$ and $X_3 = 100$, and (c) $X_2 = 10$ and $X_3 = 10$. Explain your answers.

5 Belief Propagation Implementation (25pt)

Implement the Sum-Product version of Belief Propagation in Python (using the networkx package to represent the factor graph) to compute the marginal distribution $P(A =$

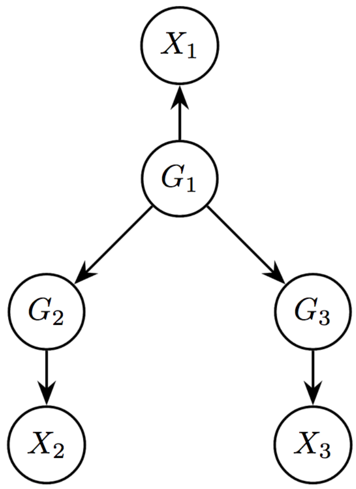

Figure 2: DPGM for Question 4.

$a, B = b)$ $\forall a, b$ for the factor graph shown in Figure 3. That is, use the graph object from networkx to define the factor graph and implement the Sum-Product algorithm to pass messages. The support of the corresponding random variables A, B, C, D and E are

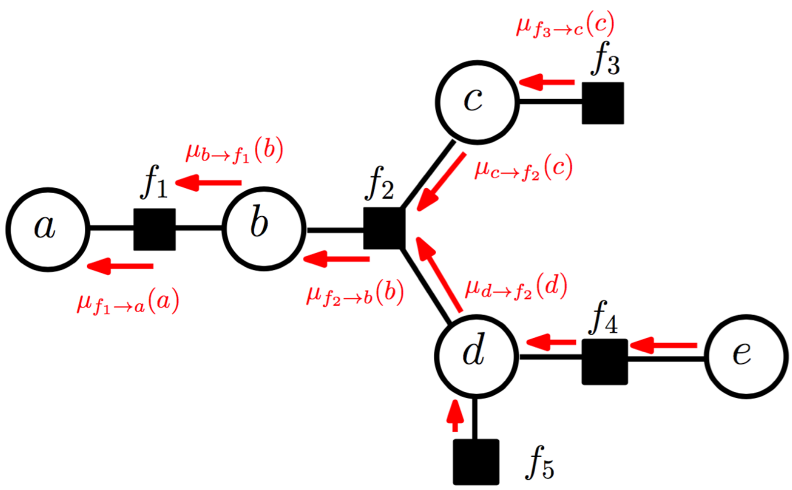

Figure 3: Factor Graph for Question 5.

$\{1, 2\}$. The factors are as follows:

1. $f_1(a, b) = a * b$ (for example, $f_1(a = 1, b = 2) = 1 * 2 = 2$).
2. $f_2(b, c, d) = 2 * (5 - b * c) - d + 1$.
3. $f_4(d, e) = d * e$.
4. $f_3(c) = 3 - c$.
5. $f_5(d) = 3 - d$.

Report the marginal distribution (plot/table) as well as your implementation (py/ipynb). Additionally,

1. Briefly describe how you implemented the algorithm (data structures and code organization).
2. What is the complexity (number of additions and multiplications if any) of computing an outgoing message from a variable node given that it is connected to F factors and has a support of k values?
3. Similarly, what is the complexity of computing an outgoing message from a factor node given it is connected to V variables each of which have a support of k values.?

6 Metropolis-Hastings Sampling (15pt)

Let $X \sim \pi$ be a distribution. To create a Monte Carlo estimate of the expectation of some function of X, i.e., $E_\pi[f(X)]$, we will do Metropolis-Hastings (MH) MCMC sampling. For this, we start with an initial $x = x_0$ and do the following:

- $x' \sim g(\cdot|x)$.
- $\alpha = \min(1, \frac{\pi(x)g(x'|x)}{pi(x')g(x|x')})$.
- With probability α accept x' as the next sample x_t.
- Set $x = x'$ and repeat.

An example of the proposal distribution is $g(x'|x) = \frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{(x'-x)^2}{\sigma^2}}$.

1. Implement the MH sampler for estimating $E[(X^2 + 10)x]$ when $\pi(x) = \frac{1}{\sqrt{2\pi}}e^{-(x-3)^2}$. Choose a suitable proposal distribution.
2. Plot the estimate as a function of the number of samples used in the estimation.

7 Gibbs Sampling (15pt)

Let the distribution $\pi(x) = \sum_{k=1}^{10} \lambda_k \mathcal{N}(x|\mu = \sqrt{k}, \sigma = 0.01k^2)$, with some arbitrarily chosen λ_k's that sum up to 1. Implement the Gibbs sampling procedure to estimate the mean. Report the mean values computed using the first 500 samples and the next 5000 samples.